

Parameter estimation for the supercritical contact process

MARTA FIOCCO¹ and WILLEM R. VAN ZWET²

¹*Department of Medical Statistics, Leiden University Medical Centre, P.O. Box 9604, 2300 RC Leiden, The Netherlands. E-mail: m.fiocco@lumc.nl*

²*Mathematical Institute, University of Leiden, P.O. Box 9512, 2300 RA Leiden, The Netherlands. E-mail: vanzwet@math.leidenuniv.nl*

Contact processes – and, more generally, interacting particle processes – can serve as models for a large variety of statistical problems, especially if we allow some simple modifications that do not essentially complicate the mathematical treatment of these processes. We begin a statistical study of the supercritical contact process that starts with a single infected site at the origin and is conditioned on survival of the infection. We consider the statistical problem of estimating the parameter λ of the process on the basis of an observation of the process at a single time t . We propose an estimator of λ and show that it is consistent and asymptotically normal as $t \rightarrow \infty$.

Keywords: contact process; parameter estimation; random mask; shrinking; supercritical

1. Introduction

A d -dimensional contact process is a simplified model for the spread of a biological organism or an infection on the lattice \mathbb{Z}^d . At each time $t \geq 0$, every point of the lattice (or site) is either infected or healthy. As time passes, a healthy site is infected at Poisson rate λ by each of its $2d$ immediate neighbours which is itself infected; an infected site recovers and becomes healthy at Poisson rate 1. Given the set of infected sites ξ_t at time t , the processes involved are independent until a change occurs. If the process starts with a set $A \subset \mathbb{Z}^d$ of infected sites at time $t = 0$, then ξ_t^A will denote the set of infected sites at time $t \geq 0$ and $\{\xi_t^A : t \geq 0\}$ will denote the contact process. For example, $\{\xi_t^{\mathbb{Z}^d} : t \geq 0\}$ or $\{\xi_t^{\{0\}} : t \geq 0\}$ will denote the processes starting with every site infected, or with a single infected site at the origin. If the starting set is chosen at random according to a probability distribution α , then the process will be written as $\{\xi_t^\alpha : t \geq 0\}$. If we do not want to specify the initial state of the process at all, we simply write $\{\xi_t : t \geq 0\}$.

We also need a compact notation for the state of a single site $x \in \mathbb{Z}^d$ at time t . For any contact process ξ_t , we write

$$\xi_t(x) = 1_{\xi_t}(x) = \begin{cases} 1 & \text{if } x \text{ is infected at time } t, \\ 0 & \text{if } x \text{ is healthy at time } t, \end{cases} \quad (1.1)$$

thus using the same symbol ξ_t for both the set of infected points and its indicator function. Of course $\xi_t^A(x)$ and $\xi_t^\alpha(x)$ will refer to the processes ξ_t^A and ξ_t^α in the same manner.

The first thing to note about the contact process is that for all non-empty $A \subset \mathbb{Z}^d$, the infection will continue forever with positive probability if and only if λ exceeds a certain *critical value* λ_d . Such a process is called *supercritical*. Thus, if we define the random hitting time

$$\tau^A = \inf\{t : \xi_t^A = \emptyset\}, \quad A \subset \mathbb{Z}^d, \quad (1.2)$$

with the convention that $\tau^A = \infty$ if $\xi_t^A \neq \emptyset$ for all $t \geq 0$, then for the supercritical contact process

$$\mathbb{P}(\tau^A = \infty) > 0 \quad (1.3)$$

for every non-empty $A \subset \mathbb{Z}^d$. Moreover, if A has infinite cardinality $|A| = \infty$, then

$$\mathbb{P}(\tau^A = \infty) = 1. \quad (1.4)$$

In the supercritical case, the process $\xi_t^{\mathbb{Z}^d}$ that starts with all sites infected converges in distribution to the so-called upper invariant measure $\nu = \nu_\lambda$. Here convergence in distribution means convergence of probabilities of events defined by the behaviour of the process on finite subsets of \mathbb{Z}^d , and ‘invariant’ refers to the fact that the process $\{\xi_t^\nu : t \geq 0\}$ is stationary. In particular, the distribution of ξ_t^ν is equal to ν for all t . Obviously, ν is also invariant under integer-valued translations of \mathbb{Z}^d . The long-range behaviour of the supercritical contact process $\{\xi_t^A : t \geq 0\}$ for arbitrary non-empty $A \subset \mathbb{Z}^d$ is described by the *complete convergence theorem*. Let μ_t^A denote the probability distribution of ξ_t^A and δ_\emptyset the distribution that assigns probability 1 to the empty set.

Theorem 1.1. *Let $A \subset \mathbb{Z}^d$ and $\lambda > \lambda_d$. Then, as $t \rightarrow \infty$,*

$$\mu_t^A \xrightarrow{w} \mathbb{P}(\tau^A < \infty)\delta_\emptyset + \mathbb{P}(\tau^A = \infty)\nu_\lambda. \quad (1.5)$$

For a proof see Liggett (1999, p. 55).

If $\lambda > \lambda_d$ and $A = \mathbb{Z}^d$, the process $\xi_t^{\mathbb{Z}^d}$ survives forever with probability 1 by (1.4) and converges exponentially to the limit process, that is, for positive C and γ and all $t \geq 0$,

$$0 \leq \mathbb{P}\left(\xi_t^{\mathbb{Z}^d}(x) = 1\right) - \mathbb{P}(\xi^\nu(x) = 1) \leq C e^{-\gamma t} \quad (1.6)$$

(Liggett 1999, p. 57).

Another major result concerning the contact process is the *shape theorem*. To formulate this result we first have to describe the graphical representation of contact processes due to Harris (1978). This is a particular coupling of all contact processes of a given dimension d and with a given value of λ , but with every possible initial state A or initial distribution α . Consider space-time $\mathbb{Z}^d \times [0, \infty)$. For every site $x \in \mathbb{Z}^d$ we define on the line $x \times [0, \infty)$ a Poisson process with rate 1; for every ordered pair (x, y) of neighbouring sites in \mathbb{Z}^d we define a Poisson process with rate λ . All of these Poisson processes are independent.

We now draw a picture of $\mathbb{Z}^d \times [0, \infty)$ where, for each site $x \in \mathbb{Z}^d$, we remove the points of the corresponding Poisson process with rate 1 from the line $x \times [0, \infty)$; for each ordered pair of neighbouring sites (x, y) we draw an arrow going perpendicularly from the

line $x \times [0, \infty)$ to the line $y \times [0, \infty)$ at the points of the Poisson processes with rate λ corresponding to the pair (x, y) .

For any set $A \subset \mathbb{Z}^d$, define ξ_t^A to be the set of sites that can be reached by starting at time 0 at some site in A and travelling until time t along unbroken segments of lines $x \times [0, \infty)$ in the direction of increasing time, as well as along arrows. Clearly, $\{\xi_t^A : t \geq 0\}$ is distributed as a contact process with initial state A . By choosing the initial set at random with distribution α , we define $\{\xi_t^\alpha : t \geq 0\}$. The obvious beauty of this coupling is that for two initial sets of infected sites $A \subset B$, we have $\xi_t^A \subset \xi_t^B$ for all $t \geq 0$.

Unless indicated otherwise, we shall assume that all contact processes are defined according to this graphical construction. We shall also restrict attention to the supercritical case and assume that $\lambda > \lambda_d$ throughout.

Before formulating the shape theorem we need to introduce some notation. Let $\|\cdot\|$ denote the L^∞ norm on \mathbb{R}^d , that is,

$$\|x\| = \max_{1 \leq i \leq d} |x_i|$$

for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, and let $Q = \{x \in \mathbb{R}^d : \|x\| \leq \frac{1}{2}\}$ denote the unit hypercube centred at the origin. For $A, B \subset \mathbb{R}^d$, $A \oplus B = \{x + y : x \in A, y \in B\}$ will denote the direct sum of A and B , and for real r , $rA = \{rx : x \in A\}$. Define

$$H_t = \bigcup_{s \leq t} \xi_s^{\{0\}} \oplus Q, \quad (1.7)$$

$$K_t = \{x \in \mathbb{Z}^d : \xi_t^{\{0\}}(x) = \xi_t^{\mathbb{Z}^d}(x)\} \oplus Q. \quad (1.8)$$

Thus for the process $\{\xi_t^{\{0\}} : t \geq 0\}$ that starts with a single infected site at the origin, H_t is obtained by taking the union of the sites that have been infected up to or at time t , and replacing these sites by unit hypercubes centred at these sites in order to fill in the space between neighbouring sites. Similarly, K_t is the filled-in version of the set of sites where $\xi_t^{\{0\}}$ and $\xi_t^{\mathbb{Z}^d}$ coincide. We are now in a position to formulate the shape theorem (cf. Durrett 1991; Bezuidenhout and Grimmett 1990).

Theorem 1.2. *There exists a bounded convex subset U of \mathbb{R}^d with the origin as an interior point and such that, for any $\epsilon \in (0, 1)$,*

$$(1 - \epsilon)tU \subset H_t \cap K_t \subset H_t \subset (1 + \epsilon)tU, \quad (1.9)$$

eventually almost surely on the event $\{\tau^{\{0\}} = \infty\}$ where $\xi_t^{\{0\}}$ survives forever.

The shape theorem describes the growth of the set of infected sites if the process $\xi_t^{\{0\}}$ survives forever. Roughly speaking, the convex hull of the set of infected sites will grow linearly in time as $t \rightarrow \infty$ and acquire an asymptotic shape tU , where U is a fixed convex set with the origin as an interior point. Inside this set, say in $(1 - \epsilon)tU$, the smallest and the largest possible process $\xi_t^{\{0\}}$ and $\xi_t^{\mathbb{Z}^d}$ are equal eventually a.s., and this must mean that, for large t , their distribution is close to the equilibrium distribution ν . Together, the complete convergence theorem and the shape theorem describe the peculiar type of convergence of the supercritical contact process to its limiting distribution. The infection spreads at a

constant speed and, relatively soon after it has reached a site x , equilibrium will set in at that site.

A third important property of the contact process is its self-duality. If, in the graphical representation, time is run backwards and all arrows representing infection of one site by another are reversed, then the new graphical representation has precisely the same probabilistic structure as the original one. In particular,

$$\mathbb{P}(\xi_t^A \cap B \neq \emptyset) = \mathbb{P}(\xi_t^B \cap A \neq \emptyset), \text{ for all } A, B \subset \mathbb{Z}^d \text{ and } t \geq 0. \quad (1.10)$$

With $A = \{0\}$ and $B = \mathbb{Z}^d$ this yields

$$\mathbb{P}(\tau^{\{0\}} > t) = \mathbb{P}(\xi_t^{\mathbb{Z}^d}(0) = 1)$$

which, letting $t \rightarrow \infty$ in the supercritical case, reduces to

$$\mathbb{P}(\tau^{\{0\}} = \infty) = \mathbb{P}(\xi_t^\nu(0) = 1).$$

Combining this with (1.6), we see that if $\lambda > \lambda_d$, then

$$\mathbb{P}(t < \tau^{\{0\}} < \infty) \leq C e^{-\gamma t} \quad (1.11)$$

(cf. Liggett 1999, p. 57).

In this paper we shall study the estimation problem for the parameter λ of the supercritical contact process $\xi_t^{\{0\}}$, given that it does not die out. Based on an observation of $\xi_t^{\{0\}}$ at a single time t , we derive an estimator $\hat{\lambda}_t^{\{0\}}$ and show that it is consistent and asymptotically normal as $t \rightarrow \infty$.

The informal description of the convergence of the contact process immediately suggests a way to derive an estimator of the parameter λ . If $\xi_t^{\{0\}}$ survives forever, then observing $\xi_t^{\{0\}}(x)$ for all sites x contained in $(1 - \epsilon)tU$ is asymptotically the same as observing the limit process $\xi_t^\nu(x)$ on this set. This asymptotic ‘equivalence’ of ξ_t^ν and $\xi_t^{\{0\}}$ on $(1 - \epsilon)tU$ should allow us to derive an estimator of λ based on the limit process $\xi_t^\nu(x)$ for sites $x \in (1 - \epsilon)tU$, and hope that this estimator will also work for the process $\xi_t^{\{0\}}$. The advantage of deriving the estimator under ξ_t^ν is that we can use the stationarity of this process to set up the estimating equation.

For $D \subset \mathbb{Z}^d$, define the total number of infected sites in the set D at time t as

$$n_t(D) = \sum_{x \in D} \xi_t(x), \quad (1.12)$$

and the total number of pairs of neighbouring sites for which one site is healthy and lies in D and the other is infected as

$$k_t(D) = \sum_{x \in D} k_t(x), \quad (1.13)$$

where

$$k_t(x) = (1 - \xi_t(x)) \sum_{|x-y|=1} \xi_t(y). \quad (1.14)$$

Here $|x - y| = \sum |x_i - y_i|$ denotes the L^1 distance between sites x and y . When we need to specify the initial state of the process we shall use an appropriate notation. For example, $n_t^{\{0\}}$ and $k_t^{\{0\}}$ will indicate that we are referring to the process $\xi_t^{\{0\}}$. Similarly, for the process ξ_t^ν , we write $n_t^\nu(x)$ and k_t^ν .

For the ξ_t^ν process, $\xi_t^\nu(x)$ increases by 1 at rate $\lambda k_t^\nu(x)$ and decreases by 1 at rate $\xi_t^\nu(x)$. As ξ_t^ν is stationary, this implies that $\lambda \mathbb{E} k_t^\nu(x) = \mathbb{E} \xi_t^\nu(x)$ and, since ξ_t^ν is spatially translation-invariant, we have

$$\lambda = \frac{\mathbb{E} \xi_t^\nu(x)}{\mathbb{E} k_t^\nu(x)} = \frac{\mathbb{E} \xi_t^\nu(0)}{\mathbb{E} k_t^\nu(0)}. \quad (1.15)$$

Notice that these expectations are independent of t because of the stationarity of ξ_t^ν . For $t \geq 0$, let $A_t \subset \mathbb{Z}^d$ be finite sets of cardinality $|A_t| \rightarrow \infty$ as $t \rightarrow \infty$. It seems reasonable to expect that some form of the law of large numbers will ensure that, as $t \rightarrow \infty$,

$$\frac{n_t^\nu(A_t)}{|A_t|} = \frac{\sum_{x \in A_t} \xi_t^\nu(x)}{|A_t|} \sim \mathbb{E} \xi_t^\nu(0)$$

and

$$\frac{k_t^\nu(A_t)}{|A_t|} = \frac{\sum_{x \in A_t} k_t^\nu(x)}{|A_t|} \sim \mathbb{E} k_t^\nu(0).$$

This would imply that $n_t^\nu(A_t)/k_t^\nu(A_t)$ is a plausible estimator of λ on the basis of an observation of the process ξ_t^ν at a single time t . If, in addition to $|A_t| \rightarrow \infty$, we also require that $A_t \subset (1 - \epsilon)tU$ for some $\epsilon > 0$, then the shape theorem suggests that, conditional on $\xi_t^{\{0\}}$ surviving forever, the probabilistic behaviour of $\xi_t^{\{0\}}$ and ξ_t^ν should be asymptotically the same on the set $A_t \subset \mathbb{Z}^d$. But this indicates that if we observe the process $\xi_t^{\{0\}}$ instead of ξ_t^ν , then $n_t^{\{0\}}(A_t)/k_t^{\{0\}}(A_t)$ would be a plausible estimator of λ based on $\xi_t^{\{0\}}$, provided that $\xi_t^{\{0\}}$ survives. Unfortunately, the set U is unknown – as is t in many applications – and hence we cannot implement this estimation procedure directly. However, the shape theorem also suggests that if $\xi_t^{\{0\}}$ survives forever, the convex hull $\mathcal{C}(\xi_t^{\{0\}})$ of the set $\xi_t^{\{0\}}$ of infected sites behaves asymptotically like tU . Hence we may expect that if we define a mask

$$C_t = (1 - \delta)\mathcal{C}(\xi_t^{\{0\}}),$$

for some $\delta > 0$, and $\xi_t^{\{0\}}$ survives, then $|C_t \cap \mathbb{Z}^d| \rightarrow \infty$ and $C_t \subset (1 - \epsilon)tU$ for some $\epsilon > 0$. Combining these ideas, we arrive at

$$\hat{\lambda}_t^{\{0\}} = \hat{\lambda}_t^{\{0\}}(C_t) = \frac{n_t^{\{0\}}(C_t)}{k_t^{\{0\}}(C_t)} \quad (1.16)$$

as a plausible estimator of λ on the basis of an observation of $\xi_t^{\{0\}}$ at a single time t . In fact we shall use masks C_t which are obtained by shrinking the set $\mathcal{C}(\xi_t^{\{0\}})$ in a more general manner than through multiplication by $1 - \delta$ (cf. Section 3).

The aim of this paper is to prove that $\hat{\lambda}_t^{\{0\}}$ is a consistent and asymptotically normal estimator of λ on the event where $\xi_t^{\{0\}}$ survives forever. To do this we not only have the considerable problem of making the above heuristic argument precise, but in order to prove

the asymptotic normality, we also have to show that, for the $\xi_t^{\{0\}}$ process conditional on survival, distant sites evolve almost independently. The technical tools for dealing with these problems are provided in Fiocco and van Zwet (2003).

We should stress at this point that shrinking $\mathcal{C}(\xi_t^{\{0\}})$ to obtain the mask C_t is absolutely essential to obtain an estimator that works well in practice. Without shrinking, the mask will contain the boundary area of the set of infected points where equilibrium has not yet set in and the infected points are therefore less dense. This has the effect of lowering the estimator of λ . Simulation shows that the resulting negative bias is considerable and that 20–40% of the sites have to be removed by shrinking to eliminate this bias (cf. Fiocco 1997). From a theoretical point of view we shall find that without shrinking – i.e. if $\delta = 0$ and hence $C_t = \mathcal{C}(\xi_t^{\{0\}})$ – we can still show consistency of the estimator $\hat{\lambda}_t^{\{0\}}$, but not its asymptotic normality.

2. Technical tools

In this section we provide the reader with a number of tools that will be used in this paper for establishing the properties of $\hat{\lambda}_t^{\{0\}}$. These results may be found in Fiocco and van Zwet (2003). Let $\mathcal{C}(\xi_t^{\{0\}})$ be the convex hull of the set of infected sites. Theorems 1.3–1.5 in Fiocco and van Zwet (2003) provide eventually almost sure bounds on this set, and probability bounds for the lower inclusion for $H_t \cap K_t$ as well as $\mathcal{C}(\xi_t^{\{0\}})$ in (1.9) and (2.1).

Theorem 2.1. *For every $\epsilon \in (0, 1)$,*

$$(1 - \epsilon)tU \subset \mathcal{C}(\xi_t^{\{0\}}) \subset (1 + \epsilon)tU \quad (2.1)$$

eventually a.s. on the set $\{\tau^{\{0\}} = \infty\}$. Moreover, for every $\epsilon \in (0, 1)$ and $r > 0$, there exists a positive number $A_{r,\epsilon}$ such that, for every $t > 0$,

$$\mathbb{P}((1 - \epsilon)tU \subset H_t \cap K_t | \tau^{\{0\}} = \infty) \geq 1 - A_{r,\epsilon} t^{-r}$$

$$\mathbb{P}((1 - \epsilon)tU \subset \mathcal{C}(\xi_t^{\{0\}}) | \tau^{\{0\}} = \infty) \geq 1 - A_{r,\epsilon} t^{-r}.$$

Before formulating the next result we need to introduce some notation. Let $H = \{0, 1\}^{\mathbb{Z}^d}$ denote the state space for the contact process. For $f : H \rightarrow \mathbb{R}$ and $x \in \mathbb{Z}^d$, define

$$\Delta_f(x) = \sup\{|f(\eta) - f(\xi)| : \eta, \xi \in H \text{ and } \eta(y) = \xi(y) \text{ for all } y \neq x\}, \quad (2.2)$$

$$\|f\| = \sum_{x \in \mathbb{Z}^d} \Delta_f(x).$$

For $R_1, R_2 \subset \mathbb{Z}^d$, let $d(R_1, R_2)$ denote the L^1 distance of R_1 and R_2 :

$$d(R_1, R_2) = \inf_{x \in R_1, y \in R_2} |x - y| = \inf_{x \in R_1, y \in R_2} \sum_{i=1}^d |x_i - y_i|.$$

Let

$$D_R = \{f : H \rightarrow \mathbb{R}, \|f\| < \infty, f(\eta) \text{ depends on } \eta \text{ only through } \eta \cap R\}, \quad (2.3)$$

that is, D_R is the class of functions f with $\|f\| < \infty$ such that $f(\eta)$ depends on η only through $\eta(x)$ with $x \in R$.

Theorem 2.2. *There exist positive numbers γ and C such that for every $R_1, R_2 \subset \mathbb{Z}^d$, $f \in D_{R_1}$, $g \in D_{R_2}$, and $t \geq 0$,*

$$\left| \text{cov} \left(f(\xi_t^{\mathbb{Z}^d}), g(\xi_t^{\mathbb{Z}^d}) \right) \right| \leq C \|f\| \cdot \|g\| e^{-\gamma d(R_1, R_2)}. \quad (2.4)$$

In particular, there exist positive numbers γ and C such that, for all $t \geq 0$, and $x, y \in \mathbb{Z}^d$,

$$\left| \text{cov} \left(\xi_t^{\mathbb{Z}^d}(x), \xi_t^{\mathbb{Z}^d}(y) \right) \right| \leq C e^{-\gamma |x-y|}, \quad (2.5)$$

and

$$\left| \text{cov} \left(k_t^{\mathbb{Z}^d}(x), k_t^{\mathbb{Z}^d}(y) \right) \right| \leq C e^{-\gamma |x-y|}. \quad (2.6)$$

Proof. The first part of the theorem is Theorem 1.7 in Fiocco and van Zwet (2003), and is proved in Section 3 of that paper. Inequalities (2.5) and (2.6) follow because $\|f\| = \|g\| = 1$ and 8, respectively. \square

Obviously (2.5) and (2.6) imply that $\sigma^2(n_t^{\mathbb{Z}^d}(D))$ and $\sigma^2(k_t^{\mathbb{Z}^d}(D))$ are of order $|D|$ for large D . The following theorem extends this results to all moments of even order.

Theorem 2.3. *For any $k = 1, 2, \dots$, there exists a number $C_k > 0$ such that for every $D \subset \mathbb{Z}^d$ and $t \geq 0$,*

$$\mu_{2k} = \mathbb{E} \left(n_t^{\mathbb{Z}^d}(D) - \mathbb{E} n_t^{\mathbb{Z}^d}(D) \right)^{2k} \leq C_k |D|^k. \quad (2.7)$$

and

$$\nu_{2k} = \mathbb{E} \left(k_t^{\mathbb{Z}^d}(D) - \mathbb{E} k_t^{\mathbb{Z}^d}(D) \right)^{2k} \leq C_k |D|^k. \quad (2.8)$$

Proof. The proof follows from Theorem 4.1 in Fiocco and van Zwet (2003). \square

Let $\tilde{\xi}_t^{\mathbb{Z}^d}$ denote a process distributed as $\xi_t^{\mathbb{Z}^d}$ conditioned on $\{\tau^{\{0\}} = \infty\}$. Theorem 1.6 in Fiocco and van Zwet (2003) asserts that we can couple the processes $\tilde{\xi}_t^{\mathbb{Z}^d}$ and $\xi_t^{\mathbb{Z}^d}$ in such a way that they coincide on tU except on a set of exponentially small probability. We shall not explicitly describe this coupling, other than to note that it is not in accordance with the graphical representation since the two processes are defined on essentially different subsets of the sample space. We repeat the theorem for the reader's convenience:

Theorem 2.4. *There exist a coupling $({}_c\xi_t^{\mathbb{Z}^d}, {}_c\tilde{\xi}_t^{\mathbb{Z}^d})$ of $(\xi_t^{\mathbb{Z}^d}, \tilde{\xi}_t^{\mathbb{Z}^d})$ and positive constants C and γ such that for all $t > 0$,*

$$\mathbb{P} \left({}_c\xi_t^{\mathbb{Z}^d} \cap tU = {}_c\tilde{\xi}_t^{\mathbb{Z}^d} \cap tU \right) > 1 - C e^{-\gamma t}.$$

Let $\tilde{\xi}_t^{\{0\}}$ denote a process which is distributed as $\xi_t^{\{0\}}$ conditioned on $\{\tau^{\{0\}} = \infty\}$. The final result in this section is a restatement of Theorem 1.8 in Fiocco and van Zwet (2003) and asserts that for this process, distant sites evolve almost independently for large t .

Theorem 2.5. *For every $\epsilon \in (0, 1)$ and $r > 0$ there exist a positive number $A_{r,\epsilon}$, as well as positive constants C and γ , such that, for all $t > 0$ and all f and g satisfying $f \in D_{R_1}$ with $R_1 \subset (1 - \epsilon)tU \cap \mathbb{Z}^d$, and $g \in D_{R_2}$ with $R_2 \subset \mathbb{Z}^d$,*

$$\left| \text{cov}\left(f(\tilde{\xi}_t^{\{0\}}), g(\tilde{\xi}_t^{\{0\}})\right) \right| \leq \|f\| \cdot \|g\| (Ce^{-\gamma d(R_2, R_2)} + A_{r,\epsilon}t^{-r}). \quad (2.9)$$

3. Shrinking

As we have argued in the Introduction, we choose the mask C_t for computing the estimator $\hat{\lambda}_t^{\{0\}}$ as a shrunken version of the convex hull $\mathcal{C}(\xi_t^{\{0\}})$ that is guaranteed to lie in $(1 - \epsilon)tU$ with large probability. As an example we discussed the choice $C_t = (1 - \delta)\mathcal{C}(\xi_t^{\{0\}})$, about which we shall have more to say later in this section (see Example 3.2). However, we also noted that it is possible to consider more general methods of shrinking, and this is the topic of the present section.

For a set $A \subset \mathbb{R}^d$ the interior of A is denoted by $\overset{\circ}{A}$ and the discrete cardinality of A as $|A|_D = |A \cap \mathbb{Z}^d|$. Define a shrinking operation as follows.

Definition 3.1. *Suppose that to any convex set $V \subset \mathbb{R}^d$ there corresponds a convex set $V^- \subset \mathbb{R}^d$. Then the map $V \rightarrow V^-$ is called a shrinking if, for every convex V and W with $0 \in \overset{\circ}{V}$,*

$$V^- \subset V, \quad (3.1)$$

$$V \subset W \Rightarrow V^- \subset W^-, \quad (3.2)$$

$$|(tV)^-|_D \rightarrow \infty \text{ as } t \rightarrow \infty, \quad (3.3)$$

and

$$\text{if } s, t \rightarrow \infty \text{ with } t/s \rightarrow 1, \text{ then } \frac{|(tV)^-|_D}{|(sV)^-|_D} \rightarrow 1. \quad (3.4)$$

Property (3.3) guarantees that if V contains a ball centred at the origin and hence tV grows linearly in t in any direction, then the number of lattice points in $(tV)^-$ tends to infinity. By a standard argument one finds that (3.4) is equivalent to the following condition: if $0 \in \overset{\circ}{V}$, then for every $\delta > 0$ there exist $\epsilon > 0$ and $t_0 > 0$ such that

$$\left| \frac{|[(1 + \epsilon)tV]^-|_D}{|[(1 - \epsilon)tV]^-|_D} - 1 \right| \leq \delta \quad \text{for all } t \geq t_0. \quad (3.5)$$

We shall base the estimator of λ on a shrunken version C_t of $\mathcal{C}(\xi_t^{\{0\}})$, that is,

$$C_t = [\mathcal{C}(\xi_t^{\{0\}})]^- \quad (3.6)$$

and

$$\hat{\lambda}_t^{\{0\}} = \hat{\lambda}_t^{\{0\}}(C_t) = \frac{n_t^{\{0\}}(C_t)}{k_t^{\{0\}}(C_t)}. \quad (3.7)$$

The set defined in (3.6) is called the random mask or window. Notice that 0 is an interior point of U and hence of $\mathcal{C}(\xi_t^{\{0\}})$ eventually a.s., so that C_t satisfies (3.1)–(3.4) eventually a.s. Since we are concerned with limit behaviour of $\xi_t^{\{0\}}$ as $t \rightarrow \infty$, this is sufficient for our purpose.

Together (3.6), (3.7) and Definition 3.1 will allow us to prove consistency of $\hat{\lambda}_t^{\{0\}}$ on the set where $\xi_t^{\{0\}}$ survives forever. However, in order to prove strong consistency of $\hat{\lambda}_t^{\{0\}}$, we need to strengthen assumption (3.3) and require that if $0 \in \mathring{V}$, then

$$\text{for some } \delta > 0, \quad \liminf_{t \rightarrow \infty} \frac{|(tV)^-|_D}{t^\delta} > 0. \quad (3.8)$$

To prove asymptotic normality of our estimator given $\{\tau^{\{0\}} = \infty\}$ we need to assume that if $0 \in \mathring{V}$, then

$$V^- \subset (1 - \delta)V, \quad (3.9)$$

while at the same time strengthening (3.3) in a different direction and requiring that

$$(tV)^- \rightarrow \mathbb{R}^d \quad \text{as } t \rightarrow \infty. \quad (3.10)$$

We end this section by presenting various ways of shrinking that one may wish to apply to the convex hull of the set of infected sites $\mathcal{C}(\xi_t^{\{0\}})$ in order to obtain the mask C_t .

Example 3.1 $V^- = V$. This satisfies Definition 3.1 as well as (3.8) and (3.10), but not (3.9). In this case we do not shrink but simply choose $C_t = \mathcal{C}(\xi_t^{\{0\}})$ for computing $\hat{\lambda}_t^{\{0\}}$.

Example 3.2 $V^- = (1 - \delta)V$, $0 < \delta < 1$. Obviously Definition 3.1 as well as (3.8)–(3.10) are satisfied. In determining the mask $C_t = (1 - \delta)\mathcal{C}(\xi_t^{\{0\}})$ we have to face the problem that we observe the set $\xi_t^{\{0\}}$, but not necessarily the location of the origin. As C_t is determined by shrinking $\mathcal{C}(\xi_t^{\{0\}})$ towards the origin, we have to estimate the origin and shrink towards this estimated origin instead. An obvious estimate of the origin is the coordinatewise average of all sites in $\mathcal{C}(\xi_t^{\{0\}})$, that is, the centre of gravity of this set of sites. In view of Theorem 2.1 and the fact that the set U is obviously symmetric with respect to the origin, it is easy to see that the estimate of the origin has error $\mathcal{O}_P(t)$ on the set where $\xi_t^{\{0\}}$ survives forever. But this implies that shrinking $\mathcal{C}(\xi_t^{\{0\}})$ towards the estimated rather than the true origin will not affect the consistency of $\hat{\lambda}_t^{\{0\}}$ in the conclusion of Theorem 4.1. The asymptotic normality of $\hat{\lambda}_t^{\{0\}}$ in Theorem 5.1 will not be affected either by a slightly more complicated argument.

Example 3.3 $V^- = \text{peeling}(V)$. This type of shrinking avoids the estimation of the origin of the picture. For an arbitrary convex set $V \subset \mathbb{R}^d$, the peeling procedure starts with the set $V_0 = \mathcal{C}(V \cap \mathbb{Z}^d)$, the convex hull of the lattice points of V . Notice that, in the particular case

we are considering, $V = \mathcal{C}(\xi_t^{\{0\}})$ and hence $V_0 = V$. The peeling of V is now obtained by removing all lattice points in the L^1 contour of V_0 , constructing the convex hull of the remaining lattice points of V_0 , and repeating this procedure k times until a fraction α of the lattice points in V_0 has been removed. This amounts to stripping away the k outermost layers of the blob. Obviously peeling satisfies Definition 3.1 as well as (3.8)–(3.10). In view of the problems encountered in Example 3.2, we prefer peeling over multiplication by $1 - \delta$ as a shrinking operation. For more details on peeling, see Fiocco (1997).

Example 3.4 $V^- = B_{\{c,r\}}$. The mask is computed by taking a Euclidean ball inside the set of infected sites with centre c and radius r , where the centre is estimated by taking the coordinatewise average of all sites in $\mathcal{C}(\xi_t^{\{0\}})$ and the radius r is computed by averaging the L^1 distances between the estimated centre and the sites in $(\mathcal{C})_{\xi_t^{\{0\}}}$.

It should be clear from these four examples that we have a great deal of freedom in choosing our mask as a shrunk version of $\mathcal{C}(\xi_t^{\{0\}})$. In order to satisfy (3.1)–(3.4), we mainly have to watch out that we do not remove all but a bounded number of lattice points of $\mathcal{C}(\xi_t^{\{0\}})$, and that for large sets the fraction α of lattice points deleted depends on the size of the set in a smooth manner. Conditions (3.8) and (3.10) are not likely to be violated for any sensible procedure either. Assumption (3.9) asserts that the shrinking is non-trivial.

Simulation of the estimator for dimension $d = 2$ indicates that for best results, the optimal fraction α of sites to be deleted by shrinking should generally be between 0.2 and 0.4, and should decrease for increasing t . For $\alpha = 0$, i.e. without shrinking, the performance of the estimator is generally disastrous. On theoretical grounds one can argue that α should be chosen proportional to t^{-1} .

4. The estimation problem: Consistency

In the proof of the consistency of $\hat{\lambda}_{t,\tau}^{\{0\}}$ we shall not follow the same route as we did in Section 1 to arrive at the estimator $\hat{\lambda}_{t,\tau}^{\{0\}}(C_t)$. Rather than introducing a new coupling to compare $\xi_t^{\{0\}}$ on $\{\tau^{\{0\}} = \infty\}$ with $\xi_{t,\tau}^\nu$ we shall simply employ the standard graphical representation for comparison with $\xi_t^{\mathbb{Z}^d}$ instead. In Theorem 2.1 we showed that on $\{\tau^{\{0\}} = \infty\}$, $\mathcal{C}(\xi_t^{\{0\}})$ can be bracketed between two non-random convex sets. By applying the shape theorem (Theorem 1.2) we reduce the problem to one concerning the $\xi_t^{\mathbb{Z}^d}$ process on a non-random convex set and then show that the difference between the random and the non-random masks is negligible.

Let $A_t \subset \mathbb{Z}^d$ be a finite non-random set with $|A_t| \rightarrow \infty$ as $t \rightarrow \infty$. By analogy with (1.12) and (1.13), define

$$n_t^{\mathbb{Z}^d}(A_t) = \sum_{x \in A_t} \xi_t^{\mathbb{Z}^d}(x) \quad (4.1)$$

$$k_t^{\mathbb{Z}^d}(A_t) = \sum_{x \in A_t} k_t^{\mathbb{Z}^d}(x), \quad k_t^{\mathbb{Z}^d}(x) = (1 - \xi_t^{\mathbb{Z}^d}(x)) \sum_{|x-y|=1} \xi_t^{\mathbb{Z}^d}(y). \quad (4.2)$$

Lemma 4.1. *Suppose that for $t \geq 0$, the sets $A_t \subset \mathbb{Z}^d$ satisfy $A_t \subset A_{t'}$ if $t < t'$, $|A_t| < \infty$ and $|A_t| \rightarrow \infty$ for $t \rightarrow \infty$. Then, as $t \rightarrow \infty$,*

$$\frac{n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} \xrightarrow{P} \mathbb{E} \xi^v(0), \quad (4.3)$$

$$\frac{k_t^{\mathbb{Z}^d}(A_t)}{|A_t|} \xrightarrow{P} \mathbb{E} k^v(0). \quad (4.4)$$

Moreover, if, for some $\delta > 0$,

$$\liminf_{t \rightarrow \infty} \frac{|A_t|}{t^\delta} > 0, \quad (4.5)$$

then, as $t \rightarrow \infty$,

$$\frac{n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} \rightarrow \mathbb{E} \xi^v(0) \text{ a.s.}, \quad (4.6)$$

$$\frac{k_t^{\mathbb{Z}^d}(A_t)}{|A_t|} \rightarrow \mathbb{E} k^v(0) \text{ a.s.} \quad (4.7)$$

Proof. We shall only prove (4.3) and (4.6). The proof of (4.4) and (4.7) is almost exactly the same.

By Theorem 2.3 and the Markov inequality,

$$\mathbb{P} \left(\left| \frac{n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} - \frac{\mathbb{E} n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} \right| \geq \epsilon \right) \leq C_{k,\epsilon} |A_t|^{-k} \quad (4.8)$$

for every $k = 1, 2, \dots$ and appropriate $C_{k,\epsilon} > 0$. By (1.6),

$$\frac{\mathbb{E} n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} = \mathbb{E} \xi_t^{\mathbb{Z}^d}(0) \rightarrow \mathbb{E} \xi^v(0) \quad (4.9)$$

as $t \rightarrow \infty$. Since $|A_t| \rightarrow \infty$, this proves (4.3).

For every $\epsilon > 0$ and $A \subset \mathbb{Z}^d$, we have

$$\mathbb{P} \left(\sup_{0 \leq s \leq h} \left| n_{t+s}^{\mathbb{Z}^d}(A) - n_t^{\mathbb{Z}^d}(A) \right| \geq \epsilon |A| \right) \leq \mathbb{P}(Z \geq \epsilon |A|), \quad (4.10)$$

where Z has a Poisson distribution with $\mathbb{E} Z = \mu = c \cdot h \cdot |A|$, where $c = 1 \vee 2d\lambda$. To see this, note that between time t and $t + h$ a change at any particular site in A occurs at rate at most c . As

$$\mathbb{E} e^Z = e^{(e-1)\mu} \leq e^{2\mu},$$

we find that if $h \leq \epsilon/(4c)$, then

$$\mathbb{P}(Z \geq \epsilon | A) \leq e^{2\mu - \epsilon |A|} \leq e^{-\epsilon |A|/2}. \quad (4.11)$$

Take $t_0 = 0$ and define $t_0 < t_1 < t_2 < \dots$ recursively by

$$t_{m+1} = (t_m + \epsilon/(4c)) \wedge \inf\{t > t_m : A_{t-} \neq A_{t+}\},$$

where

$$A_{t-} = \lim_{s \uparrow t} A_s = \bigcup_{s < t} A_s, \quad A_{t+} = \lim_{s \downarrow t} A_s = \bigcap_{s > t} A_s.$$

Hence t_{m+1} is obtained by adding to t_m until one either arrives at $t_m + \epsilon/(4c)$ or encounters a change in A_t . Because A_t is non-decreasing, this implies that by passing from t_m to t_{m+1} , one either increases t by $\epsilon/(4c)$ or $|A_t|$ by at least 1. It follows that $t_m \rightarrow \infty$ as $m \rightarrow \infty$. To see this, note that either $t_m \rightarrow \infty$ or $|A_{t_m}| \rightarrow \infty$. Since $|A_t| < \infty$ for all t , we must have $t_m \rightarrow \infty$ in both cases. Obviously there exists $0 \leq k \leq m-1$ such that $t_m \geq k\epsilon/(4c)$ and $|A_{t_{m+1}}| \geq |A_{t_m}| \geq |A_{t_{m-1}}| \geq m - k - 1$. By (4.5) this implies that

$$\liminf_m \frac{|A_{t_m-}|}{m^{\delta'}} > 0$$

for $\delta' = \delta \wedge 1$. It follows from (4.8) that, for every $k = 1, 2, \dots$,

$$\mathbb{P}\left(\left|\frac{n_{t_m}^{\mathbb{Z}^d}(A_{t_m})}{|A_{t_m}|} - \mathbb{E}\xi_{t_m}^{\mathbb{Z}^d}(0)\right| \geq \epsilon\right) \leq C'_{k,\epsilon} m^{-\delta'k}, \quad (4.12)$$

and the same is true with A_{t_m} replaced by A_{t_m-} or $A_{t_{m+}}$.

As $t_{m+1} - t_m \leq \epsilon/(4c)$ and $A_t = A_{t_{m+}}$ for $t_m < t < t_{m+1}$, (4.10) and (4.11) yield

$$\begin{aligned} \mathbb{P}\left(\sup_{t_m < t < t_{m+1}} \left|n_t^{\mathbb{Z}^d}(A_t) - n_{t_m}^{\mathbb{Z}^d}(A_{t_{m+}})\right| \geq \epsilon |A_{t_{m+}}|\right) &\leq e^{-\epsilon |A_{t_{m+}}|/2} \\ &\leq e^{-\epsilon C m^{\delta'}/2}, \end{aligned} \quad (4.13)$$

for some $C > 0$ and $m > m_0$. By (4.12) with $k > 1/\delta'$, (4.13) and the Borel–Cantelli lemma we find

$$\frac{n_t^{\mathbb{Z}^d}(A_t)}{|A_t|} - \mathbb{E}\xi_t^{\mathbb{Z}^d}(0) \rightarrow 0 \text{ a.s.,}$$

and, together with (4.9), this proves (4.6). \square

Lemma 4.1 allows us to prove both the consistency and the strong consistency of $\hat{\lambda}_t^{\{0\}}$ as $t \rightarrow \infty$.

Theorem 4.1. *Let $\hat{\lambda}_t^{\{0\}}(C_t)$ be the estimator of λ for the process $\xi_t^{\{0\}}$ defined in (3.6)–(3.7) and Definition 3.1. Then on the set where $\xi_t^{\{0\}}$ survives forever,*

$$\hat{\lambda}_t^{\{0\}}(C_t) \xrightarrow{P} \lambda \quad \text{as } t \rightarrow \infty. \quad (4.14)$$

If, in addition, (3.8) holds, then

$$\hat{\lambda}_t^{\{0\}}(C_t) \rightarrow \lambda \quad \text{as } t \rightarrow \infty \quad (4.15)$$

a.s. on the set where $\xi_t^{\{0\}}$ survives forever.

Proof. Choose $\delta > 0$ and $\epsilon > 0$ such that (3.5) is satisfied for some $t_0 > 0$. Define $A_t = [(1 - \epsilon)tU]^- \cap \mathbb{Z}^d$ and $B_t = [(1 + \epsilon)tU]^- \cap \mathbb{Z}^d$. By Theorem 2.1, (3.2) and (3.6), $A_t \subset C_t \cap \mathbb{Z}^d \subset B_t$ eventually a.s. on $\{\tau^{\{0\}} = \infty\}$, and then (3.5) ensures that

$$n_t^{\{0\}}(A_t) \leq n_t^{\{0\}}(C_t) \leq \{n_t^{\{0\}}(A_t) + |B_t \setminus A_t|\} \leq n_t^{\{0\}}(A_t) + \delta|A_t|.$$

Again by (3.5), it follows that

$$\begin{aligned} (1 + \delta)^{-1} \liminf_{t \rightarrow \infty} n_t^{\{0\}}(A_t)/|A_t| &\leq \liminf_{t \rightarrow \infty} n_t^{\{0\}}(C_t)/|C_t| \\ &\leq \limsup_{t \rightarrow \infty} n_t^{\{0\}}(C_t)/|C_t| \leq \limsup_{t \rightarrow \infty} n_t^{\{0\}}(A_t)/|A_t| + \delta. \end{aligned}$$

By (3.1)–(3.3), A_t satisfies the assumptions for (4.3) to hold and as $A_t \subset (1 - \epsilon)tU$, Theorem 1.2 implies that $n_t^{\{0\}}(A_t) = n_t^{\mathbb{Z}^d}(A_t)$ eventually a.s. on $\{\tau^{\{0\}} = \infty\}$. Letting $t \rightarrow \infty$ and then $\delta \rightarrow 0$, we find that $n_t^{\{0\}}(C_t)/|C_t| \xrightarrow{P} \mathbb{E}\xi^v(0)$ on $\{\tau^{\{0\}} = \infty\}$. In exactly the same way one may use (4.4) to prove that $k_t^{\{0\}}(C_t)/|C_t| \xrightarrow{P} \mathbb{E}k_t^v(0)$ on $\{\tau^{\{0\}} \rightarrow \infty\}$, and (4.14) follows by combining these results and using (1.15). By using (4.6) and (4.7) instead of (4.3) and (4.4), one establishes (4.15) under the additional condition (3.8). \square

Remark 4.1. By (4.14),

$$\mathbb{P}\{|\hat{\lambda}_t^{\{0\}} - \lambda| \geq \epsilon | \tau^{\{0\}} = \infty\} \rightarrow 0 \quad \text{as } t \rightarrow \infty, \quad (4.16)$$

for every $\epsilon > 0$. From a statistical point of view this appears unsatisfactory since we shall never know whether the process will survive forever and hence whether $\hat{\lambda}_t^{\{0\}}$ will be close to λ even for very large t . However, for the supercritical contact process (4.16) is obviously equivalent to

$$\mathbb{P}\{|\hat{\lambda}_t^{\{0\}} - \lambda| \geq \epsilon | \xi_t^{\{0\}} \neq \emptyset\} \rightarrow 0, \quad (4.17)$$

for every $\epsilon > 0$, and this statement does have statistical relevance. Of course our result does not provide any information in the subcritical case ($\lambda \leq \lambda_d$).

5. The estimation problem: Asymptotic normality

This section is devoted to the proof of a conditional central limit theorem for the estimator $\hat{\lambda}_t^{\{0\}} = \hat{\lambda}_t^{\{0\}}(C_t)$ based on the random mask C_t . First, we establish the joint asymptotic normality of

$$|A_t|^{-1/2} \left(n_t^{\mathbb{Z}^d}(A_t) - |A_t| \mathbb{E} \xi^v(0), k_t^{\mathbb{Z}^d}(A_t) - |A_t| \mathbb{E} k^v(0) \right)$$

for a non-random mask $A_t \subset \mathbb{Z}^d$, with $|A_t| < \infty$ for all $t \geq 0$ but $|A_t| \rightarrow \infty$ as $t \rightarrow \infty$. Next we show that this result carries over to the $\xi_t^{\{0\}}$ process, that is, the $\xi_t^{\{0\}}$ process conditioned on $\{\tau^{\{0\}} = \infty\}$. This proves the asymptotic normality of the estimator $\hat{\lambda}_t^{\{0\}}(A_t)$ given $\{\tau^{\{0\}} = \infty\}$ for a non-random mask A_t . Then we show that the contribution to the standardized estimator which is due to the randomness of the mask $C_t = [\mathcal{C}(\xi_t^{\{0\}})]^-$ vanishes as $t \rightarrow \infty$. The asymptotic normality of

$$|C_t|^{-1/2} (\hat{\lambda}_t^{\{0\}}(C_t) - \lambda)$$

given $\{\tau^{\{0\}} = \infty\}$ then follows.

A very general central limit theorem for a translation-invariant random field was proved by Bolthausen (1982) under mixing conditions. Let $\xi(x)$, $x \in \mathbb{Z}^d$, denote a real-valued translation-invariant random field, that is, $\{\xi(x) : x \in \mathbb{Z}^d\}$ is a collection of random variables and the joint law of the $\xi(x)$ is invariant under integer-valued shifts in \mathbb{Z}^d . It is assumed that $\mathbb{E} \xi^2(x) < \infty$. For $x = (x_1, \dots, x_d)$, $y = (y_1, \dots, y_d) \in \mathbb{Z}^d$, define the L^∞ distance of x and y as

$$\rho(x, y) = \max_{1 \leq i \leq d} |x_i - y_i|.$$

Let $A_n \subset \mathbb{Z}^d$, $n = 1, 2, \dots$, with $|A_n| < \infty$ for all n , $|A_n| \rightarrow \infty$ as $n \rightarrow \infty$ and

$$\frac{|\partial A_n|}{|A_n|} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (5.1)$$

Here

$$\partial A_n = \{x \in A_n : \exists y \in \mathbb{Z}^d \setminus A_n \text{ with } \rho(x, y) = 1\} \quad (5.2)$$

denotes the L^∞ contour of A_n in \mathbb{Z}^d . Consider

$$S_n = \sum_{x \in A_n} (\xi(x) - \mathbb{E} \xi(0)).$$

If $C \subset \mathbb{Z}^d$, let \mathcal{B}_C be the σ -algebra generated by $\{\xi(x), x \in C\}$. For $C_1, C_2 \subset \mathbb{Z}^d$, let

$$\rho(C_1, C_2) = \inf \{\rho(x, y) : x \in C_1, y \in C_2\}.$$

For $m \in \mathbb{N}$, $k, l \in \mathbb{N} \cup \{\infty\}$, define the mixing coefficients

$$\alpha_{k,l}(m) = \sup \{ |\mathbb{P}(B_1 \cap B_2) - \mathbb{P}(B_1)\mathbb{P}(B_2)| : B_i \in \mathcal{B}_{C_i}, |C_1| \leq k, \\ |C_2| \leq l, \rho(C_1, C_2) \geq m \}. \quad (5.3)$$

Let $N(\mu, \sigma^2)$ denote the univariate normal distribution with expectation μ and variance σ^2 and $N(\mu, \Sigma)$ the bivariate normal distribution with expectation vector μ and covariance matrix Σ . Part of Bolthausen's theorem reads as follows.

Lemma 5.1. *Suppose that, as $m \rightarrow \infty$,*

$$\sum_{m=1}^{\infty} m^{d-1} \alpha_{k,l}(m) < \infty, \quad \text{for } k + l \leq 4, \quad (5.4)$$

$$\alpha_{1,\infty}(m) = o(m^{-d}), \quad (5.5)$$

and that, for some $\delta > 0$,

$$\mathbb{E}|\zeta(x)|^{2+\delta} < \infty \quad \text{and} \quad \sum_{m=1}^{\infty} m^{d-1} \alpha_{1,1}(m)^{\delta/(2+\delta)} < \infty. \quad (5.6)$$

Then $\sum_{x \in \mathbb{Z}^d} |\text{cov}(\zeta(0), \zeta(x))| < \infty$. If, in addition, $\sigma^2 = \sum_{x \in \mathbb{Z}^d} \text{cov}(\zeta(0), \zeta(x)) > 0$ and (5.1) holds, then $|A_n|^{-1/2} S_n / \sigma$ converges in distribution to $N(0, 1)$.

For our purposes we have to modify this result slightly. First of all, we allow a different stationary random field $\zeta_n(x)$ for each n , so that S_n becomes

$$\tilde{S}_n = \sum_{x \in A_n} (\zeta_n(x) - \mathbb{E}\zeta_n(0)).$$

As a result, we also have to replace the assumptions of the lemma by versions which are uniform in n . This means that in the assumptions of the lemma we replace $\alpha_{k,l}(m)$ by the supremum over n of expression (5.3) for $\zeta_n(x)$. Similarly, the integrability of $|\zeta_n(x)|^{2+\delta}$ in (5.6) is replaced by the uniform integrability of $|\zeta_n(x)|^{2+\delta}$. Then Bolthausen's proof goes through to show that $\sup_n \sum_{x \in \mathbb{Z}^d} |\text{cov}(\zeta_n(0), \zeta_n(x))| < \infty$ and that $|A_n|^{-1/2} \tilde{S}_n / \sigma_n \xrightarrow{\mathcal{D}} N(0, 1)$, provided that $\liminf \sigma_n^2 > 0$, where $\sigma_n^2 = \sum_{x \in \mathbb{Z}^d} \text{cov}(\zeta_n(0), \zeta_n(x))$.

A second modification of Lemma 5.1 concerns assumption (5.5). It is clear from Bolthausen's proof that (5.5) may be replaced by

$$\alpha_{1,l}(l^{1/(2d+1)}) = o(l^{-1/2}) \quad \text{as } l \rightarrow \infty. \quad (5.7)$$

With these modifications, Lemma 5.1 allows us to prove:

Lemma 5.2. Choose $\epsilon \in (0, 1)$ and $A_t \subset \mathbb{Z}^d$ for $t \geq 0$ such that

$$A_t \subset (1 - \epsilon)tU, \quad |A_t| \rightarrow \infty, \quad \text{and} \quad |\partial A_t|/|A_t| \rightarrow 0 \quad \text{as } t \rightarrow \infty. \quad (5.8)$$

As $t \rightarrow \infty$, the conditional distribution of the random vector

$$|A_t|^{-1/2} \left(\sum_{x \in A_t} (\xi_t^{\{0\}}(x) - \mathbb{E}\xi^v(0)), \sum_{x \in A_t} (k_t^{\{0\}}(x) - \mathbb{E}k^v(0)) \right) \quad (5.9)$$

given $\{\tau^{\{0\}} = \infty\}$ converges weakly to $N(0, \Sigma)$, where

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix} \quad (5.10)$$

and

$$\sigma_1^2 = \sum_{x \in \mathbb{Z}^d} \text{cov}(\xi^\nu(0), \xi^\nu(x)), \quad \sigma_2^2 = \sum_{x \in \mathbb{Z}^d} \text{cov}(k^\nu(0), k^\nu(x)), \quad \sigma_{1,2} = \sum_{x \in \mathbb{Z}^d} \text{cov}(k^\nu(0), \xi^\nu(x)). \quad (5.11)$$

Proof. The lemma concerns the process $\bar{\xi}_t^{\{0\}}$ which is distributed as $\xi_t^{\{0\}}$ conditioned on $\{\tau^{\{0\}} = \infty\}$, restricted to the set $(1 - \epsilon)tU$. By Theorems 1.2 and 2.4 we may first replace this process by the conditional process $\bar{\xi}_t^{\mathbb{Z}^d}$ and then by the unconditional process $\xi_t^{\mathbb{Z}^d}$. Similarly, we may replace $\mathbb{E}\xi^\nu(0)$ by $\mathbb{E}\xi_t^{\mathbb{Z}^d}(0)$ since $|A_t|^{1/2}|\mathbb{E}\xi_t^{\mathbb{Z}^d}(0) - \mathbb{E}\xi^\nu(0)| = \mathcal{O}(t^{d/2}e^{-\gamma t}) \rightarrow 0$ by (1.6). The same holds for $\mathbb{E}k^\nu(0)$ and $\mathbb{E}k_t^{\mathbb{Z}^d}(0)$. Hence, it suffices to prove that

$$|A_t|^{-1/2} \left(\sum_{x \in A_t} \left(\xi_t^{\mathbb{Z}^d}(x) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(0) \right), \sum_{x \in A_t} \left(k_t^{\mathbb{Z}^d}(x) - \mathbb{E}k_t^{\mathbb{Z}^d}(0) \right) \right).$$

is asymptotically $N(0, \Sigma)$.

Let u and v be real numbers and define

$$\zeta_t(x) = u\xi_t^{\mathbb{Z}^d}(x) + vk_t^{\mathbb{Z}^d}(x).$$

Clearly $\{\zeta_t(x), x \in \mathbb{Z}^d\}$ is a real-valued, translation-invariant random field for each t . Consider

$$\tilde{S}_t = \sum_{x \in A_t} (\zeta_t(x) - \mathbb{E}\zeta_t(0)).$$

The fact that \tilde{S}_t depends on a real-valued index $t \rightarrow \infty$, instead of an integer $n \rightarrow \infty$ as in our version of Bolthausen's result, is of course immaterial in what follows. Note that $|\zeta_t(x)| \leq |u| + 4|v|$ so that all moments of $|\zeta_t(x)|$ are bounded independent of t .

Let us write $\alpha_{kl}(m)$ for the quantity defined in (5.3) computed for ζ_t . By Theorem 2.2 and because $\rho(x, y) \leq d(x, y) = \sum_{i=1}^d |x_i - y_i|$, there exist positive C and γ such that

$$\alpha_{kl}(m) \leq Ckle^{-\gamma m},$$

independent of t . This means that assumptions (5.4), (5.6) and (5.7) are satisfied uniformly in t . Note that (5.5) is not satisfied since we cannot allow $l = \infty$, but, as we have indicated, (5.7) serves just as well. Hence, we have proved that

$$|A_t|^{-1/2} \sigma_t^{-1} \sum_{x \in A_t} \left(u \left(\xi_t^{\mathbb{Z}^d}(x) - \mathbb{E}\xi_t^{\mathbb{Z}^d}(0) \right) + v \left(k_t^{\mathbb{Z}^d}(x) - \mathbb{E}k_t^{\mathbb{Z}^d}(0) \right) \right) \quad (5.12)$$

has a standard normal limit distribution provided that $\liminf \sigma_t^2 > 0$. Here

$$\sigma_t^2 = \sum_{x \in \mathbb{Z}^d} \text{cov} \left(u\xi_t^{\mathbb{Z}^d}(0) + vk_t^{\mathbb{Z}^d}(0), u\xi_t^{\mathbb{Z}^d}(x) + vk_t^{\mathbb{Z}^d}(x) \right). \quad (5.13)$$

By (1.6) the terms in (5.13) converge to $\text{cov}(u\xi^\nu(0) + vk_t^\nu(0), u\xi^\nu(x) + vk_t^\nu(x))$ as $t \rightarrow \infty$, and by Theorem 2.2 the terms are bounded by $C' \exp\{-\gamma \sum_{1 \leq i \leq d} |x_i|\}$, independent of t . It follows that the sum also converges, so σ_t^2 tends to

$$\sigma^2(u, v) = \sum_{x \in \mathbb{Z}^d} \text{cov}(u\xi^v(0) + vk_t^v(0), u\xi^v(x) + vk_t^v(x)). \quad (5.14)$$

Hence \tilde{S}_t is asymptotically $N(0, \sigma^2(u, v))$ if $\sigma^2(u, v) > 0$ and asymptotically degenerate at 0 if $\sigma^2(u, v) = 0$. The lemma is proved by the Cramér–Wold device. \square

To prove the joint asymptotic normality of $n_t^{\{0\}}(C_t)$ and $k_t^{\{0\}}(C_t)$ – and hence of $\hat{\lambda}_t^{\{0\}}(C_t)$ – conditional on $\{\tau^{\{0\}} = \infty\}$, we have to consider the difference between these quantities computed for the random mask C_t and a non-random mask which is close to C_t . For $\epsilon > 0$ and $t > 0$, define

$$A_t = [(1 - \epsilon)tU]^- \cap \mathbb{Z}^d, \quad B_t = [(1 + \epsilon)tU]^- \cap \mathbb{Z}^d, \quad (5.15)$$

that is, A_t and B_t consist of the sites in the shrunken versions of the sets $(1 - \epsilon)tU$ and $(1 + \epsilon)tU$ respectively, where the shrinking operation $V \rightarrow V^-$ is defined in Definition 3.1.

Lemma 5.3. *For $\epsilon \in (0, 1)$ define A_t and B_t as in (5.15) and let $D_t = (B_t \setminus A_t) \cap C_t$, with $C_t = [\mathcal{C}(\xi_t^{\{0\}})]^-$ as given by (3.6) and Definition 3.1. If the shrinking operation $V \rightarrow V^-$ satisfies (3.9) for some $\delta \in (0, 1)$, then, for every $z > 0$,*

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in D_t} (\xi_t^{\{0\}}(x) - \mathbb{E}\xi^v(0)) \right| \geq z | \tau^{\{0\}} = \infty \right) = 0, \quad (5.16)$$

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in D_t} (k_t^{\{0\}}(x) - \mathbb{E}k^v(0)) \right| \geq z | \tau^{\{0\}} = \infty \right) = 0. \quad (5.17)$$

Proof. We shall only prove (5.16) as the proof of (5.17) is almost the same. As before, we write $\tilde{\xi}_t^{\{0\}}$ for the conditional process $(\xi_t^{\{0\}} | \tau^{\{0\}} = \infty)$; \mathbb{P} will denote the conditional probability $\mathbb{P}(\cdot | \tau^{\{0\}} = \infty)$.

Without loss of generality we assume that $\epsilon \leq \delta/4$ so that $(1 - \delta)(1 + \epsilon) \leq 1 - 3\delta/4$ and, by (3.9),

$$B_t = [(1 + \epsilon)tU]^- \cap \mathbb{Z}^d \subset (1 - \delta)(1 + \epsilon)tU \subset \left(1 - \frac{3\delta}{4}\right)tU. \quad (5.18)$$

As $|D_t| \leq |B_t| = \mathcal{O}(t^d)$ and $|A_t| \rightarrow \infty$, we note that in (5.16) we may replace $\mathbb{E}\xi^v(0)$ first by $\mathbb{E}\xi_t^{\mathbb{Z}^d}(0)$ because of (1.6) and then by $\mathbb{E}\tilde{\xi}_t^{\mathbb{Z}^d}(0)$ because of Theorem 2.4, and finally by $\mathbb{E}\tilde{\xi}_t^{\{0\}}(0)$ in view of Theorem 2.1. Hence, in order to prove (5.23), it is enough to show that

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in D_t} (\tilde{\xi}_t^{\{0\}}(x) - \mathbb{E}\tilde{\xi}_t^{\{0\}}(x)) \right| \geq z \right) = 0. \quad (5.19)$$

Define

$$C_t^* = \left[\mathcal{C}(\{\tilde{\xi}_t^{\{0\}} \cup (1 - \epsilon)tU\} \cap (1 + \epsilon)tU) \right]^-. \quad (5.20)$$

By (2.1), $(1 - \epsilon)tU \subset \mathcal{C}(\tilde{\xi}_t^{\{0\}}) \subset (1 + \epsilon)tU$, and hence

$$\begin{aligned}\mathcal{C}\left(\{\bar{\xi}_t^{\{0\}} \cup (1-\epsilon)tU\} \cap (1+\epsilon)tU\right) &= \mathcal{C}\left(\bar{\xi}_t^{\{0\}} \cup (1-\epsilon)tU\right) \\ &= \mathcal{C}\left(\mathcal{C}(\bar{\xi}_t^{\{0\}} \cup (1-\epsilon)tU)\right) = \mathcal{C}\left(\bar{\xi}_t^{\{0\}}\right)\end{aligned}$$

eventually a.s. $(\bar{\mathbb{P}})$. It follows that

$$C_t^* = [\mathcal{C}(\bar{\xi}_t^{\{0\}})]^- = C_t \quad (5.21)$$

eventually a.s. $(\bar{\mathbb{P}})$. Obviously this implies that

$$\begin{aligned}\sum_{x \in D_t} \left(\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x) \right) &= \sum_{x \in B_t \setminus A_t} \left(\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x) \right) I_{C_t}(x) \\ &= \sum_{x \in B_t \setminus A_t} \left(\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x) \right) I_{C_t^*}(x)\end{aligned}$$

eventually a.s. $(\bar{\mathbb{P}})$. Instead of (5.19), it is therefore sufficient to show that

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in B_t \setminus A_t} (\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x)) I_{C_t^*}(x) \right| \geq z \right) = 0.$$

Clearly this will follow if we prove that

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} |A_t|^{-1} \mathbb{E} \left[\sum_{x \in B_t \setminus A_t} (\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x)) I_{C_t^*}(x) \right]^2 = 0. \quad (5.22)$$

By (5.20) the random set C_t^* is determined by the random set $\{\bar{\xi}_t^{\{0\}} \cup (1-\epsilon)tU\} \cap (1+\epsilon)tU$ which is bracketed by the non-random convex sets $(1-\epsilon)tU$ and $(1+\epsilon)tU$. It follows that C_t^* is determined by the values of $\bar{\xi}_t^{\{0\}}(y)$ for sites $y \in (1+\epsilon)tU \setminus (1-\epsilon)tU$. Put differently, for every $x \in \mathbb{Z}^d$, the function $g_x : H \rightarrow \{0, 1\}$ defined by

$$g_x(\bar{\xi}_t^{\{0\}}) = I_{C_t^*}(x) \quad (5.23)$$

satisfies

$$g_x \in D_R, \quad \text{with } R = \{(1+\epsilon)tU \setminus (1-\epsilon)tU\} \cap \mathbb{Z}^d \quad (5.24)$$

and D_R defined by (2.3).

The expected value in (5.22) can be written as

$$\begin{aligned}
 & \mathbb{E} \left[\sum_{x \in B_t \setminus A_t} \left(\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x) \right) I_{C_t^*}(x) \right]^2 \\
 &= \sum_{x, x' \in B_t \setminus A_t} \mathbb{E} \left(\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x) \right) \left(\bar{\xi}_t^{\{0\}}(x') - \mathbb{E} \bar{\xi}_t^{\{0\}}(x') \right) I_{C_t^*}(x) I_{C_t^*}(x') \quad (5.25) \\
 &= \sum_{x, x' \in B_t \setminus A_t} \mathbb{E} f_x(\bar{\xi}_t^{\{0\}}) f_{x'}(\bar{\xi}_t^{\{0\}}) g_x(\bar{\xi}_t^{\{0\}}) g_{x'}(\bar{\xi}_t^{\{0\}}),
 \end{aligned}$$

with $f_x(\bar{\xi}_t^{\{0\}}) = \bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x)$ and g_x defined by (5.23). Obviously

$$f_x \cdot f_{x'} \in D_{\{x, x'\}}, \quad g_x \cdot g_{x'} \in D_R, \quad (5.26)$$

in view of (5.24). If $x, x' \in B_t \setminus A_t$, then (5.18) ensures that $\{x, x'\} \subset (1 - 3\delta/4)tU$ and, because $\epsilon \leq \delta/4$, (5.24) implies that $R \subset \{(1 - \delta/4)tU\}^c$. Hence, if $d(\cdot, \cdot)$ denotes L^1 distance, then

$$d(\{x, x'\}, R) \geq b_\delta''' t \quad \text{for all } x, x' \in B_t \setminus A_t, \quad (5.27)$$

where b_δ''' is a positive number depending only on δ . Finally, we use (2.2) to compute

$$\|f_x \cdot f_{x'}\| = 2, \quad \|g_x \cdot g_{x'}\| \leq |R| \leq a\epsilon t^d \leq a\delta t^d, \quad (5.28)$$

for an appropriate constant $a > 0$. Combining (5.25)–(5.28) and invoking Theorem 2.5 with $r = 3d$, we obtain

$$\begin{aligned}
 & |A_t|^{-1} \mathbb{E} \left[\sum_{x \in B_t \setminus A_t} (\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \bar{\xi}_t^{\{0\}}(x)) I_{C_t^*}(x) \right]^2 \\
 & \leq |A_t|^{-1} \sum_{x, x' \in B_t \setminus A_t} \mathbb{E} f_x(\bar{\xi}_t^{\{0\}}) f_{x'}(\bar{\xi}_t^{\{0\}}) \mathbb{E} g_x(\bar{\xi}_t^{\{0\}}) g_{x'}(\bar{\xi}_t^{\{0\}}) + M_t \\
 & \leq |A_t|^{-1} \sum_{x, x' \in B_t \setminus A_t} |\text{cov}(\bar{\xi}_t^{\{0\}}(x), \bar{\xi}_t^{\{0\}}(x'))| + M_t,
 \end{aligned}$$

where the remainder term M_t satisfies, for appropriate positive c_δ and c'_δ ,

$$\begin{aligned}
 |M_t| & \leq |A_t|^{-1} |B_t \setminus A_t|^2 c_\delta \|f_x \cdot f_{x'}\| \cdot \|g_x \cdot g_{x'}\| t^{-3d} \\
 & \leq c'_\delta |A_t|^{-1} \rightarrow 0 \quad \text{as } t \rightarrow \infty,
 \end{aligned}$$

since $|B_t \setminus A_t| \leq |B_t| \leq |(1 + \epsilon)tU|_D \leq |(1 + \delta/4)tU|_D = \mathcal{O}(t^d)$ and $|A_t| \rightarrow \infty$ by (3.3).

To prove (5.22), it therefore remains to be shown that

$$\lim_{c \rightarrow 0} \limsup_{t \rightarrow \infty} |A_t|^{-1} \sum_{x, x' \in B_t \setminus A_t} \left| \text{cov}(\bar{\xi}_t^{\{0\}}(x), \bar{\xi}_t^{\{0\}}(x')) \right| = 0. \quad (5.29)$$

Invoking Theorem 2.5 once more, this time with $r = d + 1$, we find that, for $x, x' \in B_t \setminus A_t$, $x \neq x'$, and appropriate $c_\delta'' > 0$,

$$\left| \text{cov} \left(\bar{\xi}_t^{\{0\}}(x), \bar{\xi}_t^{\{0\}}(x') \right) \right| \leq c''_d |x - x'|^{-(d+1)},$$

since $x, x' \in B_t \setminus A_t$ implies $|x - x'| = \mathcal{O}(t)$. It follows that

$$\begin{aligned} \sum_{x, x' \in B_t \setminus A_t} \left| \text{cov} \left(\bar{\xi}_t^{\{0\}}(x), \bar{\xi}_t^{\{0\}}(x') \right) \right| &\leq |B_t \setminus A_t| \left(1 + c''_d \sum_{x \in \mathbb{Z}^d, x \neq 0} |x|^{-(d+1)} \right) \\ &\leq c'''_d |B_t \setminus A_t| \end{aligned}$$

for some $c'''_d > 0$, as $\sum_{x \in \mathbb{Z}^d \setminus \{0\}} |x|^{-(d+1)}$ converges. Hence, (5.29) holds if

$$\lim_{\epsilon \rightarrow 0} \limsup_{t \rightarrow \infty} \frac{|B_t \setminus A_t|}{|A_t|} = 0.$$

But since $A_t = [(1 - \epsilon)tU]^- \cap \mathbb{Z}^d$ and $B_t = [(1 + \epsilon)tU]^- \cap \mathbb{Z}^d$, this is a consequence of (3.5). This proves (5.29) and the lemma. \square

We are now in a position to prove the main result of this paper.

Theorem 5.1. *Let $\hat{\lambda}_t^{\{0\}}(C_t)$ be the estimator of λ for the process $\xi_t^{\{0\}}$ defined in (3.6)–(3.7) and Definition 3.1. If the shrinking operation $V \rightarrow V^-$ satisfies (3.9) for some $\delta \in (0, 1)$ as well as (3.10), then, as $t \rightarrow \infty$, the conditional distribution of*

$$|C_t|_D^{1/2} [\hat{\lambda}_t^{\{0\}}(C_t) - \lambda], \quad (5.30)$$

given that $\{\tau^{\{0\}} = \infty\}$, converges weakly to $N(0, \sigma^2)$. Here

$$\sigma^2 = \lambda^2 \left[\frac{\sigma_1^2}{\{\mathbb{E} \xi^v(0)\}^2} + \frac{\sigma_2^2}{\{\mathbb{E} k^v(0)\}^2} - \frac{2\sigma_{1,2}}{\{\mathbb{E} \xi^v(0) \mathbb{E} k^v(0)\}} \right], \quad (5.31)$$

where σ_1 , σ_2 and $\sigma_{1,2}$ are given by (5.11).

Proof. In the proof we write $\bar{\xi}_t^{\{0\}}$ for the conditional process $(\xi_t^{\{0\}} | \tau^{\{0\}} = \infty)$. For $t \geq 0$, define A_t and B_t by (5.15). Since $0 \in \mathring{U}$, we have $A_t \subset (1 - \epsilon)tU$ by (3.1) and $[(1 - \epsilon)tU]^- \rightarrow \mathbb{R}^d$ by (3.10). Because $[(1 - \epsilon)tU]^-$ is bounded and convex, it follows that $|\partial A_t|/|A_t| \rightarrow 0$ as $t \rightarrow \infty$ by an easy argument. Hence A_t satisfies condition (5.8) of Lemma 5.2 and we find that, for every $\epsilon \in (0, 1)$, the random vector

$$\left(|A_t|^{-1/2} \left[\sum_{x \in A_t} (\bar{\xi}_t^{\{0\}}(x) - \mathbb{E} \xi^v(0)) \right], |A_t|^{-1/2} \left[\sum_{x \in A_t} (\bar{k}_t^{\{0\}}(x) - \mathbb{E} k^v(0)) \right] \right), \quad (5.32)$$

has a limiting $N(0, \Sigma)$ distribution with Σ given by (5.10)–(5.11).

In view of (3.9), we may apply Lemma 5.3 to obtain

$$\begin{aligned}\Psi_1(\epsilon) &= \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in D_t} (\xi_t^{\{0\}}(x) - \mathbb{E} \xi^\nu(0)) \right| \geq z \right) \rightarrow 0, \\ \Psi_1(\epsilon) &= \limsup_{t \rightarrow \infty} \mathbb{P} \left(|A_t|^{-1/2} \left| \sum_{x \in D_t} (\bar{k}_t^{\{0\}}(x) - \mathbb{E} k^\nu(0)) \right| \geq z \right) \rightarrow 0,\end{aligned}\tag{5.33}$$

as $\epsilon \rightarrow 0$ for every $z > 0$. Here $D_t = (B_t \setminus A_t) \cap C_t$. Notice that by (2.1) and (3.2) we have $A_t \subset C_t \cap \mathbb{Z}^d \subset B_t$ and hence

$$C_t \cap \mathbb{Z}^d = A_t \cup D_t, \quad A_t \cap D_t = \emptyset,\tag{5.34}$$

eventually a.s. on $\{\tau^{\{0\}} = \infty\}$.

Next we note that $|B_t|/|A_t| \rightarrow 1$ by (3.5) and hence $|C_t|/|A_t| \rightarrow 1$ eventually a.s. on $\{\tau^{\{0\}} = \infty\}$ as $t \rightarrow \infty$. It now follows by a standard argument that the limit distribution of (5.32) will remain unchanged if A_t is replaced by C_t and $|A_t|$ by $|C_t|$. Finally, (3.7) and another standard argument establish the theorem. \square

6. The asymptotic variance of $\hat{\lambda}_t^{\{0\}}(C_t)$

If the variance σ^2 of the normal limit distribution in Theorem 5.1 were known, then this would allow us to assess the accuracy of the estimator or to set up asymptotic confidence intervals for λ of the form

$$\hat{\lambda}_t^{\{0\}}(C_t) - u_{\alpha/2} |C_t|_D^{-1/2} \sigma < \lambda < \hat{\lambda}_t^{\{0\}}(C_t) + u_{\alpha/2} |C_t|_D^{-1/2} \sigma,\tag{6.1}$$

where u_α is the upper α -point of the standard normal distribution. This asymptotic confidence interval would be valid provided that $\xi_t^{\{0\}}$ survives forever, but, as we pointed out in Remark 4.1, it is enough that $\xi_t^{\{0\}} \neq \emptyset$, that is, that the process has survived up to time t .

Since σ^2 is unknown we have to find an estimator of σ^2 . One way to achieve this would be to estimate $\sigma^2 = \sigma^2(\lambda)$ as a function of λ by simulating $\xi_t^{\{0\}}$ a large number of times for each λ , each time computing the value of $\hat{\lambda}_t^{\{0\}}(C_t)$ and using $|C_t|_D$ times the sample variance of these values as an estimate of $\sigma^2(\lambda)$. One could then use $\sigma^2(\hat{\lambda}_t^{\{0\}}(C_t))$ as an estimate of σ^2 . Of course in any particular instance it would be enough to carry out these simulations only for $\lambda = \hat{\lambda}_t^{\{0\}}(C_t)$.

An alternative way to estimate σ^2 would be to use the observed process $\xi_t^{\{0\}}$ itself. First, we subdivide the mask C_t into k subsets $C_{t,1}, \dots, C_{t,k}$ of (approximately) equal size and compute the values $\hat{\lambda}_t^{\{0\}}(C_{t,i})$ for $i = 1, \dots, k$. We then use $k^{-1}|C_t|_D$ times the sample variance of these values as an estimate of σ^2 .

An obvious advantage of the second method is that it is not as dependent on the model as the first. It is quite conceivable that the estimator $\hat{\lambda}_t^{\{0\}}(C_t)$ is a useful statistic in a much broader class of models than the contact process. In this case the second method is more likely to produce a sensible result than the first.

Acknowledgements

The authors are indebted to an associate editor and two referees for constructive criticism. They also wish to thank the past editor, Ole Barndorff-Nielsen, for his willingness to continue handling this paper well after the end of his tenure.

References

- Bezuidenhout, C. and Grimmett, G. (1990) The critical contact process dies out. *Ann. Probab.*, **18**, 1462–1482.
- Bolthausen, E. (1982) On the central limit theorem for stationary mixing random fields. *Ann. Probab.*, **10**, 1047–1050.
- Durrett, R. (1991) The contact process, 1974–1989. In W.E. Kohler and B.S. White (eds), *Mathematics of Random Media*, Lectures in Appl. Math. 27, pp. 1–18. Providence, RI: American Mathematical Society.
- Fiocco, M. (1997) Statistical estimation for the supercritical contact process. Doctoral thesis, University of Leiden, The Netherlands.
- Fiocco, M. and van Zwet, W.R. (2003) Decaying correlations for the supercritical contact process conditioned on survival. *Bernoulli*, **9**, 763–781.
- Harris, T.E. (1978) Additive set-valued Markov processes and graphical methods. *Ann. Probab.*, **6**, 355–378.
- Liggett, T. (1999) *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. New York: Springer-Verlag.

Received April 1999 and revised April 2003